

=====

Sequence Listing was accepted.

If you need help call the Patent Electronic Business Center at (866) 217-9197 (toll free).

Reviewer: markspencer

Timestamp: [year=2008; month=9; day=17; hr=15; min=6; sec=54; ms=28;]

=====

Application No: 10574554 Version No: 2.0

Input Set:

Output Set:

Started: 2008-08-18 14:55:36.874
Finished: 2008-08-18 14:55:39.767
Elapsed: 0 hr(s) 0 min(s) 2 sec(s) 893 ms
Total Warnings: 21
Total Errors: 0
No. of SeqIDs Defined: 21
Actual SeqID Count: 21

Error code	Error Description
W 402	Undefined organism found in <213> in SEQ ID (1)
W 402	Undefined organism found in <213> in SEQ ID (2)
W 402	Undefined organism found in <213> in SEQ ID (3)
W 402	Undefined organism found in <213> in SEQ ID (4)
W 402	Undefined organism found in <213> in SEQ ID (5)
W 402	Undefined organism found in <213> in SEQ ID (6)
W 402	Undefined organism found in <213> in SEQ ID (7)
W 402	Undefined organism found in <213> in SEQ ID (8)
W 402	Undefined organism found in <213> in SEQ ID (9)
W 402	Undefined organism found in <213> in SEQ ID (10)
W 213	Artificial or Unknown found in <213> in SEQ ID (11)
W 213	Artificial or Unknown found in <213> in SEQ ID (12)
W 213	Artificial or Unknown found in <213> in SEQ ID (13)
W 213	Artificial or Unknown found in <213> in SEQ ID (14)
W 213	Artificial or Unknown found in <213> in SEQ ID (15)
W 213	Artificial or Unknown found in <213> in SEQ ID (16)
W 213	Artificial or Unknown found in <213> in SEQ ID (17)
W 213	Artificial or Unknown found in <213> in SEQ ID (18)
W 213	Artificial or Unknown found in <213> in SEQ ID (19)
W 213	Artificial or Unknown found in <213> in SEQ ID (20)

Input Set:

Output Set:

Started: 2008-08-18 14:55:36.874
Finished: 2008-08-18 14:55:39.767
Elapsed: 0 hr(s) 0 min(s) 2 sec(s) 893 ms
Total Warnings: 21
Total Errors: 0
No. of SeqIDs Defined: 21
Actual SeqID Count: 21

Error code	Error Description
W 213	Artificial or Unknown found in <213> in SEQ ID (21)

SEQUENCE LISTING

<110> De Maria, Leonardo
Andersen, Carsten
Christensen, Lars Lehmann Hylling
Lassen, Soren Flensted
Ostergaard, Peter Rahbek

<120> Protease Variants

<130> 10508.204-US

<140> 10574554
<141> 2008-08-18

<150> DK PCT/DK2004/000688
<151> 2004-10-08

<150> DK PA 2003 01494
<151> 2003-10-10

<150> DK PA 2004 00333
<151> 2004-03-01

<150> US 60/549,347
<151> 2004-03-02

<150> US 60/510,450
<151> 2003-10-10

<160> 21

<170> PatentIn version 3.5

<210> 1
<211> 1596
<212> DNA
<213> Nocardiopsis sp. NRRL 18262 ("Protease 10")

<220>
<221> CDS
<222> (318)..(1463)

<220>
<221> sig_peptide
<222> (318)..(404)

<220>
<221> mat_peptide
<222> (900)..(1463)

<400> 1
acgtttggta cgggtaccgg tgtccgcatt tggccagaat gcccccttgc gacagggAAC 60

ggattcggtc ggttagcgcat cgactccgac aaccgcgagg tggccgttcg cgtcgccacg 120

ttctgcacc gtcatgcac ccatcatcggttgcaccccgccagctctgatggccacc 180
gttctgacgg tcttccctc accaaaacgt gcacctatgg ttaggacgtt gttaccgaa 240
tgtctcggtg aacgacaggg gcccggacggattcggcccc gatccccgt tgatcccccc 300
aggagagtag ggacccc atg cga ccc tcc ccc gtt gtc tcc gcc atc ggt 350
Met Arg Pro Ser Pro Val Val Ser Ala Ile Gly
-190 -185

acg gga gcg ctg gcc ttc ggt ctg gcg ctg tcc ggt acc ccg ggt 395
Thr Gly Ala Leu Ala Phe Gly Leu Ala Leu Ser Gly Thr Pro Gly
-180 -175 -170

gcc ctc gcg gcc acc gga gcg ctc ccc cag tca ccc acc ccg gag 440
Ala Leu Ala Ala Thr Gly Ala Leu Pro Gln Ser Pro Thr Pro Glu
-165 -160 -155

gcc gac gcg gtc tcc atg cag gag gcg ctc cag cgc gac ctc gac 485
Ala Asp Ala Val Ser Met Gln Glu Ala Leu Gln Arg Asp Leu Asp
-150 -145 -140

ctg acc tcc gcc gag gcc gag gag ctg ctg gcc gcc cag gac acc 530
Leu Thr Ser Ala Glu Ala Glu Glu Leu Leu Ala Ala Gln Asp Thr
-135 -130 -125

gcc ttc gag gtc gac gag gcc gcg gcc gag gcc gcc ggg gac gcc 575
Ala Phe Glu Val Asp Glu Ala Ala Ala Glu Ala Ala Gly Asp Ala
-120 -115 -110

tat ggc ggc tcc gtc ttc gac acc gag agc ctg gaa ctg acc gtc ctg 623
Tyr Gly Gly Ser Val Phe Asp Thr Glu Ser Leu Glu Leu Thr Val Leu
-105 -100 -95

gtc acc gat gcc gcc gcg gtc gag gcc gtg gag gcc acc ggc gcc ggg 671
Val Thr Asp Ala Ala Ala Val Glu Ala Val Glu Ala Thr Gly Ala Gly
-90 -85 -80

acc gag ctg gtc tcc tac ggc atc gac ggt ctc gac gag atc gtc cag 719
Thr Glu Leu Val Ser Tyr Gly Ile Asp Gly Leu Asp Glu Ile Val Gln
-75 -70 -65

gag ctc aac gcc gcc gac gcc gtt ccc ggt gtg gtc ggc tgg tac ccg 767
Glu Leu Asn Ala Ala Asp Ala Val Pro Gly Val Val Gly Trp Tyr Pro
-60 -55 -50 -45

gac gtg gcg ggt gac acc gtc gtc ctg gag gtc ctg gag ggt tcc gga 815
Asp Val Ala Gly Asp Thr Val Val Leu Glu Val Leu Glu Gly Ser Gly
-40 -35 -30

gcc gac gtc agc ggc ctg ctc gcg gac gcc ggc gtg gac gcc tcg gcc 863
Ala Asp Val Ser Gly Leu Leu Ala Asp Ala Gly Val Asp Ala Ser Ala
-25 -20 -15

gtc gag gtg acc acg agc gac cag ccc gag ctc tac gcc gac atc atc 911
Val Glu Val Thr Thr Ser Asp Gln Pro Glu Leu Tyr Ala Asp Ile Ile
-10 -5 -1 1

ggt ggt ctg gcc tac acc atg ggc ggc cgc tgt tcg gtc ggc ttc gcg	959
Gly Gly Leu Ala Tyr Thr Met Gly Gly Arg Cys Ser Val Gly Phe Ala	
5 10 15 20	
gcc acc aac gcc gcc ggt cag ccc ggg ttc gtc acc gcc ggt cac tgc	1007
Ala Thr Asn Ala Ala Gly Gln Pro Gly Phe Val Thr Ala Gly His Cys	
25 30 35	
ggc cgc gtg ggc acc cag gtg acc atc ggc aac ggc agg ggc gtc ttc	1055
Gly Arg Val Gly Thr Gln Val Thr Ile Gly Asn Gly Arg Gly Val Phe	
40 45 50	
gag cag tcc gtc ttc ccc ggc aac gac gcg gcc ttc gtc cgc ggt acg	1103
Glu Gln Ser Val Phe Pro Gly Asn Asp Ala Ala Phe Val Arg Gly Thr	
55 60 65	
tcc aac ttc acg ctg acc aac ctg gtc acg cgc tac aac acc ggc ggg	1151
Ser Asn Phe Thr Leu Thr Asn Leu Val Ser Arg Tyr Asn Thr Gly Gly	
70 75 80	
tac gcc acg gtc gcc ggt cac aac cag gcc ccc atc ggc tcc tcc gtc	1199
Tyr Ala Thr Val Ala Gly His Asn Gln Ala Pro Ile Gly Ser Ser Val	
85 90 95 100	
tgc cgc tcc ggc tcc acc acc ggt tgg cac tgc ggc acc atc cag gcc	1247
Cys Arg Ser Gly Ser Thr Thr Gly Trp His Cys Gly Thr Ile Gln Ala	
105 110 115	
cgc ggc cag tcg gtg agc tac ccc gag ggc acc gtc acc aac atg acc	1295
Arg Gly Gln Ser Val Ser Tyr Pro Glu Gly Thr Val Thr Asn Met Thr	
120 125 130	
cgg acc acc gtg tgc gcc gag ccc ggc gac tcc ggc ggc tcc tac atc	1343
Arg Thr Thr Val Cys Ala Glu Pro Gly Asp Ser Gly Ser Tyr Ile	
135 140 145	
tcc ggc acc cag gcc cag ggc gtg acc tcc ggc ggc tcc ggc aac tgc	1391
Ser Gly Thr Gln Ala Gln Gly Val Thr Ser Gly Gly Ser Gly Asn Cys	
150 155 160	
cgc acc ggc ggg acc acc ttc tac cag gag gtc acc ccc atg gtg aac	1439
Arg Thr Gly Thr Thr Phe Tyr Gln Glu Val Thr Pro Met Val Asn	
165 170 175 180	
tcc tgg ggc gtc cgt ctc cgg acc tgatccccgc ggttccaggc ggaccgacgg	1493
Ser Trp Gly Val Arg Leu Arg Thr	
185	
tcgtgacctg agtaccaggc gtccccggc cttccagcgg cgtccgcacc ggggtgggac	1553
cgggcgtggc cacggccccca cccgtgaccg gaccgccccgg cta	1596

<210> 2
<211> 382
<212> PRT

<213> Nocardiopsis sp. NRRL 18262 ("Protease 10")

<400> 2

Met Arg Pro Ser Pro Val Val Ser Ala Ile Gly Thr Gly Ala Leu
-190 -185 -180

Ala Phe Gly Leu Ala Leu Ser Gly Thr Pro Gly Ala Leu Ala Ala
-175 -170 -165

Thr Gly Ala Leu Pro Gln Ser Pro Thr Pro Glu Ala Asp Ala Val
-160 -155 -150

Ser Met Gln Glu Ala Leu Gln Arg Asp Leu Asp Leu Thr Ser Ala
-145 -140 -135

Glu Ala Glu Glu Leu Leu Ala Ala Gln Asp Thr Ala Phe Glu Val
-130 -125 -120

Asp Glu Ala Ala Ala Glu Ala Ala Gly Asp Ala Tyr Gly Gly Ser
-115 -110 -105

Val Phe Asp Thr Glu Ser Leu Glu Leu Thr Val Leu Val Thr Asp Ala
-100 -95 -90

Ala Ala Val Glu Ala Val Glu Ala Thr Gly Ala Gly Thr Glu Leu Val
-85 -80 -75

Ser Tyr Gly Ile Asp Gly Leu Asp Glu Ile Val Gln Glu Leu Asn Ala
-70 -65 -60

Ala Asp Ala Val Pro Gly Val Val Gly Trp Tyr Pro Asp Val Ala Gly
-55 -50 -45

Asp Thr Val Val Leu Glu Val Leu Glu Gly Ser Gly Ala Asp Val Ser
-40 -35 -30 -25

Gly Leu Leu Ala Asp Ala Gly Val Asp Ala Ser Ala Val Glu Val Thr
-20 -15 -10

Thr Ser Asp Gln Pro Glu Leu Tyr Ala Asp Ile Ile Gly Gly Leu Ala
-5 -1 1 5

Tyr Thr Met Gly Gly Arg Cys Ser Val Gly Phe Ala Ala Thr Asn Ala

10 15 20

Ala Gly Gln Pro Gly Phe Val Thr Ala Gly His Cys Gly Arg Val Gly
25 30 35 40

Thr Gln Val Thr Ile Gly Asn Gly Arg Gly Val Phe Glu Gln Ser Val
45 50 55

Phe Pro Gly Asn Asp Ala Ala Phe Val Arg Gly Thr Ser Asn Phe Thr
60 65 70

Leu Thr Asn Leu Val Ser Arg Tyr Asn Thr Gly Gly Tyr Ala Thr Val
75 80 85

Ala Gly His Asn Gln Ala Pro Ile Gly Ser Ser Val Cys Arg Ser Gly
90 95 100

Ser Thr Thr Gly Trp His Cys Gly Thr Ile Gln Ala Arg Gly Gln Ser
105 110 115 120

Val Ser Tyr Pro Glu Gly Thr Val Thr Asn Met Thr Arg Thr Thr Val
125 130 135

Cys Ala Glu Pro Gly Asp Ser Gly Ser Tyr Ile Ser Gly Thr Gln
140 145 150

Ala Gln Gly Val Thr Ser Gly Gly Ser Gly Asn Cys Arg Thr Gly Gly
155 160 165

Thr Thr Phe Tyr Gln Glu Val Thr Pro Met Val Asn Ser Trp Gly Val
170 175 180

Arg Leu Arg Thr
185

<210> 3
<211> 1065
<212> DNA
<213> Nocardiopsis dassonvillei subspecies dassonvillei DSM 43235 ("Protease 18")

<220>
<221> CDS
<222> (1)...(1062)

<220>
<221> mat_peptide
<222> (499)..(1062)

<400> 3
gct ccg gcc ccc gtc ccc cag acc ccc gtc gcc gac gac agc gcc 45
Ala Pro Ala Pro Val Pro Gln Thr Pro Val Ala Asp Asp Ser Ala
-165 -160 -155

gcc agc atg acc gag gcg ctc aag cgc gac ctc gac ctc acc tcg 90
Ala Ser Met Thr Glu Ala Leu Lys Arg Asp Leu Asp Leu Thr Ser
-150 -145 -140

gcc gag gcc gag gag ctt ctc tcg gcg cag gaa gcc gcc atc gag 135
Ala Glu Ala Glu Glu Leu Leu Ser Ala Gln Glu Ala Ala Ile Glu
-135 -130 -125

acc gac gcc gag gcc acc gag gcc gcg ggc gag gcc tac ggc ggc 180
Thr Asp Ala Glu Ala Thr Glu Ala Ala Gly Glu Ala Tyr Gly Gly
-120 -115 -110

tca ctg ttc gac acc gag acc ctc gaa ctc acc gtg ctg gtc acc gac 228
Ser Leu Phe Asp Thr Glu Thr Leu Glu Leu Thr Val Leu Val Thr Asp
-105 -100 -95

gcc tcc gcc gtc gag gcg gtc gag gcc acc gga gcc cag gcc acc gtc 276
Ala Ser Ala Val Glu Ala Val Glu Ala Thr Gly Ala Gln Ala Thr Val
-90 -85 -80 -75

gtc tcc cac ggc acc gag ggc ctg acc gag gtc gtg gag gac ctc aac 324
Val Ser His Gly Thr Glu Gly Leu Thr Glu Val Val Glu Asp Leu Asn
-70 -65 -60

ggc gcc gag gtt ccc gag agc gtc ctc ggc tgg tac ccg gac gtg gag 372
Gly Ala Glu Val Pro Glu Ser Val Leu Gly Trp Tyr Pro Asp Val Glu
-55 -50 -45

agc gac acc gtc gtg gtc gag gtg ctg gag ggc tcc gac gcc gac gtc 420
Ser Asp Thr Val Val Glu Val Leu Glu Gly Ser Asp Ala Asp Val
-40 -35 -30

gcc gcc ctg ctc gcc gac gcc ggt gtg gac tcc tcc tcg gtc cgg gtg 468
Ala Ala Leu Leu Ala Asp Ala Gly Val Asp Ser Ser Ser Val Arg Val
-25 -20 -15

gag gag gcc gag gag gcc ccg cag gtc tac gcc gac atc atc ggc ggc 516
Glu Glu Ala Glu Glu Ala Pro Gln Val Tyr Ala Asp Ile Ile Gly Gly
-10 -5 -1 1 5

ctg gcc tac tac atg ggc ggc cgc tgc tcc gtc ggc ttc gcc gcg acc 564
Leu Ala Tyr Tyr Met Gly Gly Arg Cys Ser Val Gly Phe Ala Ala Thr
10 15 20

aac agc gcc ggt cag ccc ggt ttc gtc acc gcc ggc cac tgc ggc acc 612
Asn Ser Ala Gly Gln Pro Gly Phe Val Thr Ala Gly His Cys Gly Thr
25 30 35

gtc ggc acc ggc gtg acc atc ggc aac ggc acc ggc acc ttc cag aac	660
Val Gly Thr Gly Val Thr Ile Gly Asn Gly Thr Gly Thr Phe Gln Asn	
40 45 50	
tcg gtc ttc ccc ggc aac gac gcc gcc ttc gtc cgc ggc acc tcc aac	708
Ser Val Phe Pro Gly Asn Asp Ala Ala Phe Val Arg Gly Thr Ser Asn	
55 60 65 70	
ttc acc ctg acc aac ctg gtc tcg cgc tac aac tcc ggc ggc tac cag	756
Phe Thr Leu Thr Asn Leu Val Ser Arg Tyr Asn Ser Gly Gly Tyr Gln	
75 80 85	
tcg gtg acc ggt acc agc cag gcc ccg gcc ggc tcg gcc gtg tgc cgc	804
Ser Val Thr Gly Thr Ser Gln Ala Pro Ala Gly Ser Ala Val Cys Arg	
90 95 100	
tcc ggc tcc acc acc ggc tgg cac tgc ggc acc atc cag gcc cgc aac	852
Ser Gly Ser Thr Thr Gly Trp His Cys Gly Thr Ile Gln Ala Arg Asn	
105 110 115	
cag acc gtg cgc tac ccg cag ggc acc gtc tac tcg ctc acc cgc acc	900
Gln Thr Val Arg Tyr Pro Gln Gly Thr Val Tyr Ser Leu Thr Arg Thr	
120 125 130	
aac gtg tgc gcc gag ccc ggc gac tcc ggc ggt tcg ttc atc tcc ggc	948
Asn Val Cys Ala Glu Pro Gly Asp Ser Gly Gly Ser Phe Ile Ser Gly	
135 140 145 150	
tcg cag gcc cag ggc gtc acc tcc ggc tcc ggc aac tgc tcc gtc	996
Ser Gln Ala Gln Gly Val Thr Ser Gly Ser Gly Asn Cys Ser Val	
155 160 165	
ggc ggc acg acc tac tac cag gag gtc acc ccg atg atc aac tcc tgg	1044
Gly Gly Thr Thr Tyr Tyr Gln Glu Val Thr Pro Met Ile Asn Ser Trp	
170 175 180	
ggt gtc agg atc cgg acc taa	1065
Gly Val Arg Ile Arg Thr	
185	

<210> 4
<211> 354
<212> PRT
<213> Nocardiopsis dassonvillei subspecies dassonvillei DSM 43235 ("Protease 18")

<400> 4

Ala Pro Ala Pro Val Pro Gln Thr Pro Val Ala Asp Asp Ser Ala
-165 -160 -155

Ala Ser Met Thr Glu Ala Leu Lys Arg Asp Leu Asp Leu Thr Ser
-150 -145 -140

Ala Glu Ala Glu Glu Leu Leu Ser Ala Gln Glu Ala Ala Ile Glu

-135	-130	-125
Thr Asp Ala Glu Ala Thr Glu Ala Ala Gly Glu Ala Tyr Gly Gly		
-120	-115	-110
Ser Leu Phe Asp Thr Glu Thr Leu Glu Leu Thr Val Leu Val Thr Asp		
-105	-100	-95
Ala Ser Ala Val Glu Ala Val Glu Ala Thr Gly Ala Gln Ala Thr Val		
-90	-85	-80
-75		
Val Ser His Gly Thr Glu Gly Leu Thr Glu Val Val Glu Asp Leu Asn		
-70	-65	-60
Gly Ala Glu Val Pro Glu Ser Val Leu Gly Trp Tyr Pro Asp Val Glu		
-55	-50	-45
Ser Asp Thr Val Val Val Glu Val Leu Glu Gly Ser Asp Ala Asp Val		
-40	-35	-30
Ala Ala Leu Leu Ala Asp Ala Gly Val Asp Ser Ser Ser Val Arg Val		
-25	-20	-15
Glu Glu Ala Glu Glu Ala Pro Gln Val Tyr Ala Asp Ile Ile Gly Gly		
-10	-5	-1 1
5		
Leu Ala Tyr Tyr Met Gly Gly Arg Cys Ser Val Gly Phe Ala Ala Thr		
10	15	20
Asn Ser Ala Gly Gln Pro Gly Phe Val Thr Ala Gly His Cys Gly Thr		
25	30	35
Val Gly Thr Gly Val Thr Ile Gly Asn Gly Thr Gly Thr Phe Gl		